



**LISBOA  
SCHOOL OF  
ECONOMICS &  
MANAGEMENT**

**MESTRADO  
MATEMÁTICA FINANCEIRA**

**TRABALHO FINAL DE MESTRADO  
RELATÓRIO**

**APLICAÇÃO DA ESTATÍSTICA BAYESIANA AO  
RISCO DE CRÉDITO**

**ADRIANO DINIS OLIVEIRA**

**SETEMBRO - 2014**



**LISBOA  
SCHOOL OF  
ECONOMICS &  
MANAGEMENT**

**MESTRADO  
MATEMÁTICA FINANCEIRA**

**TRABALHO FINAL DE MESTRADO  
RELATÓRIO**

**APLICAÇÃO DA ESTATÍSTICA BAYESIANA AO  
RISCO DE CRÉDITO**

**ADRIANO DINIS OLIVEIRA**

**ORIENTAÇÃO:**

**NOME: DOUTOR FILIPE CORREIA PACCETTI**

**NOME: PROFESSOR DOUTOR JOÃO AFONSO BASTOS**

**SETEMBRO - 2014**

## **Agradecimentos**

Em primeiro lugar gostaria de agradecer à Deloitte, em especial ao Dr. Filipe Pacetti, pela disponibilidade e apoio que sempre demonstrou para esclarecer todas as minhas dúvidas.

Agradeço também ao meu orientador académico, Professor Doutor João Afonso Bastos, pela disponibilidade demonstrada no desenvolvimento deste trabalho.

Por fim, gostaria de agradecer à minha família, pai, mãe, irmão e irmã, que sempre estiveram ao meu lado. Sem eles seria impossível.

## Resumo

O cálculo de probabilidade de incumprimento de uma carteira de crédito é essencial para o cálculo dos requisitos de fundos mínimos dos bancos. No entanto, existem diversas circunstâncias em que a informação bancária não é suficiente, ou fiável, fazendo com que uma análise baseada apenas em dados históricos não seja apropriada.

O objetivo principal deste projeto é o de desenvolver e implementar um modelo que seja capaz de incorporar a informação fornecida por um perito com a informação histórica de uma carteira de crédito.

Para atingir o objetivo recorreu-se à estatística Bayesiana que permite incorporar, de forma coerente, as duas fontes de informação.

O estudo recai sobre uma carteira de crédito de empresas de um banco português.

Os resultados do projeto apontam para que o valor médio da probabilidade de incumprimento da função *a priori* e *a posteriori* sejam semelhantes, no entanto a função *a posteriori* tem uma menor dispersão. É também constatado que existe uma correlação temporal positiva apesar de não ser muito forte.

**Palavras-chave:** Risco de Crédito, Estatística Bayesiana, Monte Carlo via cadeias de Markov

## Abstract

The calculation of probability of default of a loan portfolio is essential for computing the minimum capital requirements that banks need to keep. However, there are several circumstances in which bank data is scarce or not reliable, making historical data analysis not appropriate.

The main goal of this project is to develop and implement a model able to incorporate information given by an expert and historical data.

To pursue the main goal, we have used Bayesian statistics which allows coherent incorporation of both kinds of information (subjective and objective).

This study analyses a commercial loan portfolio of a portuguese bank.

The results indicate that the expected probability of default given by the expert is similar to the expected probability of default computed using the posteriori function. However, the *posteriori* function has lower dispersion than the *priori*. It is also found a weak positive correlation between time periods.

**Keywords:** Credit Risk, Bayesian Statistics, Markov chain Monte Carlo

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Risco de Crédito</b>	<b>5</b>
<b>3</b>	<b>Estatística Bayesiana</b>	<b>8</b>
3.1	Teorema de Bayes . . . . .	8
3.2	Função <i>a priori</i> . . . . .	9
3.2.1	Recolha de Informação . . . . .	10
3.3	Monte Carlo via Cadeias de Markov . . . . .	12
3.3.1	Cadeias de Markov . . . . .	12
3.3.2	Algoritmo Metropolis-Hastings . . . . .	16
3.3.3	Algoritmo de Gibbs . . . . .	19
3.3.4	Algoritmo de Metropolis-Hastings dentro de Gibbs . . . . .	19
3.3.5	Análise de resultados das cadeias de Markov . . . . .	20
<b>4</b>	<b>Metodologia e Resultados</b>	<b>23</b>
4.1	Dados . . . . .	23
4.2	Modelo . . . . .	25
4.2.1	Função <i>a priori</i> . . . . .	27

Adriano Oliveira	Aplicação da Estatística Bayesiana ao Risco de Crédito	vii
4.2.2	Função <i>a posteriori</i> . . . . .	29
4.2.3	Implementação e Resultados . . . . .	30
<b>5</b>	<b>Conclusões</b>	<b>34</b>
	<b>Referências Bibliográficas</b>	<b>39</b>
<b>A</b>	<b>Esquema Clássico e Bayesiano</b>	<b>40</b>
<b>B</b>	<b>Distribuição <i>a priori</i></b>	<b>41</b>
<b>C</b>	<b>Algoritmo M-H dentro de Gibbs</b>	<b>42</b>
<b>D</b>	<b>Diagnóstico Metropolis-Hastings</b>	<b>43</b>
D.1	Convergência . . . . .	43
D.1.1	Médias e Teste Geweke . . . . .	43
D.1.2	Teste Raftery e Lewis . . . . .	44
D.2	Correlação . . . . .	44

# Lista de Tabelas

4.1	Dados da carteira Grandes Empresas . . . . .	24
4.2	Médias das variáveis <i>a posteriori</i> . . . . .	32
D.1	Teste Geweke . . . . .	43
D.2	Teste Raftery e Lewis . . . . .	44



# Lista de Figuras

2.1	Função densidade de probabilidade da perda . . . . .	6
4.1	Distribuição <i>a priori</i> . . . . .	28
4.2	Função $\pi(p \mathbf{y})$ . . . . .	32
4.3	Função $\pi(\tau \mathbf{y})$ . . . . .	33
A.1	Esquema Clássico (Paulino et al. 2003) . . . . .	40
A.2	Esquema Bayesiano (Paulino et al. 2003) . . . . .	40
B.1	Desfasamento dos quantis da distribuição <i>a priori</i> beta e dos quantis do especialista . . . . .	41
D.1	Evolução da Média da variável $p$ . . . . .	43
D.2	Evolução da Média da variável $\tau$ . . . . .	43
D.3	Correlação dos resultados . . . . .	44
D.4	Correlação depois do método Batch . . . . .	45

# Lista de Algoritmos

1	Metropolis-Hastings . . . . .	17
2	Gibbs . . . . .	19
3	Metropolis-Hastings dentro de Gibbs . . . . .	20
4	Metropolis-Hastings dentro de Gibbs Utilizado . . . . .	42

# Abreviaturas

**EAD** valor da exposição dado incumprimento.

**EL** perda esperada.

**IDE** ambiente de desenvolvimento integrado.

**IRB** métodos dos ratings internos.

**LGD** perda dado incumprimento.

**MCMC** Monte Carlo Via cadeias de Markov.

**MCSE** erro standard de Monte Carlo.

**PD** probabilidade de incumprimento.

**UL** perda inesperada.

# 1. Introdução

Nas últimas duas décadas a gestão de risco tem vindo a ser alvo de uma crescente preocupação, tanto por parte dos bancos como por parte da supervisão.

Em 1988, é apresentado o primeiro acordo de Basileia, que tinha como objetivos a harmonização global da supervisão bancária em diversos países bem como fortalecer a estabilidade e solidez do sistema bancário. O acordo focava-se, principalmente, no risco de crédito e tinha como principal preocupação garantir que o rácio entre o capital elegível e os ativos ponderados pelo risco fosse superior ou igual a 8% .

No entanto, devido à metodologia simples de mensuração do risco descrita no acordo e à inovação do sistema financeiro e da tecnológica, que ocorreu principalmente ao longo da década de 90, tornou-se necessário atualizar o acordo de Basileia I. É assim que surge, em 2004, a versão final do Basileia II e em 2010 o acordo de Basileia III.

O acordo de Basileia III, à semelhança do segundo acordo, baseia-se em três pilares: Pilar I - Requisitos mínimos de capital; Pilar II - Processo de Supervisão; Pilar III - Disciplina de Mercado. O estudo realizado centra-se no Pilar I, mais precisamente na área de risco de crédito, sendo as restantes áreas o risco de mercado e operacional.

O acordo propõem dois tipos de modalidades de mensuração do risco de crédito: A abordagem "*Padrão*" e a abordagem de métodos dos ratings internos (IRB) que por sua vez se subdivide em duas: "*Foundation*" e "*Advanced*".

A abordagem Padrão está sobretudo dependente das cotações de rating e de ponderadores de risco definidos pelo Comité de Basileia.

A abordagem IRB pretende promover o desenvolvimento, pelos bancos, de metodologias de mensuração do risco. No entanto, estas estão sujeitas à validação do regulador.

Desde que foi introduzida, a abordagem IRB que as carteiras de crédito com um histórico de poucos incumprimentos têm sido alvo de estudo devido à dificuldade de calcular a probabilidade de incumprimento (PD). De facto, na problemática da mensuração do risco de carteiras com poucos incumprimentos, levou mesmo à divulgação de uma carta explicativa pelo Comité de Basileia de Supervisão Bancária, em resposta às preocupações levantadas pelos bancos.

A principal preocupação demonstrada era com a falta de dados de incumprimento, o que levava a que tais carteiras fossem excluídos da análise avançada (Basel Committee on Banking Supervision 2005b).

Apesar de a carta não sugerir especificamente a abordagem Bayesiana, propõe que se utilizem outras fontes de informação. No entanto, a abordagem Bayesiana parece ser a melhor opção para resolver o problema, uma vez que permite a inserção de opiniões de peritos no cálculo das probabilidades de incumprimento (Kiefer 2011).

O objetivo do estágio, realizado na Deloitte Consultores, é o de desenvolver e implementar uma metodologia de cálculo das probabilidades de incumprimento para carteiras que tenham registado poucos incumprimentos.

As razões para a utilização da estatística Bayesiana não se esgotam na falta

de informação de dados, uma vez que também pode ser utilizada quando: o custo de recolha de informação é demasiado elevado ou quando a qualidade dos dados é questionável (Jacobs & Kiefer 2010).

O problema de investigação proposto reside no cálculo de probabilidades de incumprimento através da combinação da informação fornecida por dados históricos com a informação de um perito.

O estudo realizado tem por base dados de uma carteira de empréstimos a empresas com um rating elevado de um banco português, que irá ser mencionado por Banco A. O nome do perito também será omitido, sendo referido como Perito A.

Tendo por base o problema de investigação, neste trabalho irão ser estudadas metodologias de eliciação da opinião de peritos, algo a que não tem sido dado suficiente atenção na literatura (Jacobs & Kiefer 2010). A eliciação é um processo pelo qual se formula uma função de distribuição que represente o conhecimento e expectativas de uma pessoa sobre um dado evento aleatório. Os processos de eliciação são bastante frequentes em projetos que envolvam a estatística Bayesiana para criar as funções de distribuição *a priori* (Garthwaite et al. 2005).

O relatório está estruturado da seguinte forma. No capítulo 2 é feita uma introdução ao risco de crédito. No capítulo 3 apresenta-se a estatística Bayesiana, metodologias de eliciação e, fundamentalmente, Monte Carlo via cadeias de Markov. De seguida, é descrita a metodologia utilizada para o cálculo da probabilidade de incumprimento e a sua implementação, bem como os dados utilizados

no estudo. Neste mesmo capítulo são também apresentados os resultados. Por fim, serão apresentadas as conclusões do relatório.

## 2. Risco de Crédito

Como seria de esperar, o crédito representa grande parte dos ativos de um banco. Em 2013, o crédito representava aproximadamente 68% dos ativos do sistema bancário português, valor que tem vindo a decrescer desde 2010 (Banco de Portugal 2014).

A gestão do risco de crédito é inerente à atividade bancária e tem o objetivo de gerir a possibilidade de o devedor não cumprir com as obrigações estipuladas no contrato, maximizando a taxa de retorno e mantendo uma exposição ao risco de crédito dentro de parâmetros razoáveis (Basel Committee on Banking Supervision 2000).

Um aspeto essencial na análise de risco de crédito é a definição de incumprimento. Segundo Basel Committee on Banking Supervision (2006), o incumprimento ocorre quando uma das duas situações seguintes se verifica:

1. existe um atraso no pagamento superior a 90 dias, ou
2. o banco reconhece que o devedor não tem qualquer intenção de cumprir com as suas obrigações.

É então de extrema importância que os bancos utilizem e desenvolvam metodologias de mensuração de risco de crédito adequadas a cada devedor ou a cada carteira. Na atividade bancária é sempre incerto o volume das perdas dos bancos, no entanto espera-se que os bancos sejam capazes de calcular uma previsão anual da perda esperada (EL). Tais perdas devem ser provisiona-



das pelos bancos. No entanto, para além da perda esperada, existe também a perda inesperada (UL) que se espera que não aconteça todos os anos, mas quando acontece tem um grande impacto no banco. Tendo como objetivo proteger os bancos de perdas que possam por em risco a sua existência surge, assim, legislação com base nos acordos de Basileia. É imposto o cálculo de requisitos de capital que sirva para cobrir essas perdas inesperadas.

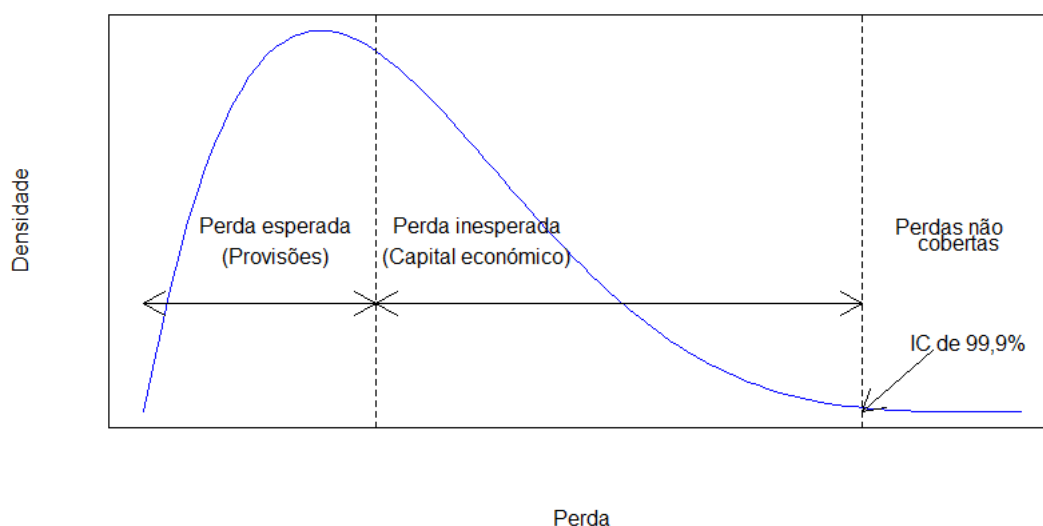


Figura 2.1: Função densidade de probabilidade da perda

O risco de crédito foi alvo de grande preocupação no acordo de Basileia II, uma vez que a metodologia proposta por Basileia I, em 2004, já se encontrava ultrapassada.

O acordo de Basileia II e III não diferem nas metodologias apresentadas para o risco de crédito. A metodologia *Standard* é baseada em notações de rating atribuídas pelas agências. A diferenciação entre a ponderação de risco a atribuir aos créditos depende da área de atuação do devedor e do tipo de exposição.

A abordagem IRB promove o desenvolvimento por parte dos bancos de metodologias de mensuração do risco e divide-se em duas: *Foundation* e *Advanced*.

A abordagem foca-se em três fatores:

1. A probabilidade de incumprimento de um dado devedor (PD);
2. A perda ocorrida dado que existe incumprimento (perda dado incumprimento (LGD));
3. O valor da exposição dado incumprimento (EAD).

Na abordagem *Foundation* os bancos podem, havendo aprovação do supervisor nacional, calcular a PD enquanto que os outros parâmetros são fixados externamente. Na abordagem *Advanced* os bancos, obtendo aprovação, podem calcular para além da PD, também a LGD.

É natural que os bancos optem por desenvolver metodologias internas, uma vez que, normalmente, estas levam a que os requisitos de capital sejam menos elevados em relação à abordagem *Standard*. Tal deve-se à existência de um maior critério no cálculo (Antão & Lacerda 2008).

## 3. Estatística Bayesiana

Em estatística existem duas grandes filosofias. A mais conhecida é a abordagem frequentista. Esta abordagem interpreta o conceito de probabilidade através da frequência relativa de um acontecimento, quando a amostra tende para infinito, ou o simples rácio entre os casos favoráveis e os casos possíveis. Esta abordagem tem como objetivo utilizar a informação recolhida através de experiências passadas para determinar um, ou vários, parâmetros desconhecidos, que são tidos como constantes.

A abordagem Bayesiana toma como variáveis aleatórias os parâmetros a estimar, uma vez que se desconhece seu valor. Assim, o conceito de probabilidade é interpretado como grau de credibilidade. Passa a ser necessário definir uma distribuição para o parâmetro (função *a priori*) uma vez que o parâmetro é incerto e, segundo a estatística Bayesiana, toda a incerteza deve ser quantificada em termos de probabilidades. A definição da função *a priori* introduz assim uma componente subjetiva ao problema. No anexo A está representada uma esquema lógico que representa as diferenças entre as duas abordagens.

### 3.1 Teorema de Bayes

O teorema de Bayes assume um papel fundamental uma vez que é a partir dele que se combina a função de verosimilhança, que contém a informação da amostra, com a informação *a priori* do parâmetro que é agora uma variável aleatória.

**Teorema de Bayes 1.** Sendo  $x$  uma realização de  $X$ , a distribuição a posteriori, com  $\theta \in \Theta$ , é dada por:

$$f(\theta|x) = \frac{f(x|\theta) f(\theta)}{\int_{\Theta} f(x|\theta) f(\theta) d\theta} \quad (3.1)$$

Observando o numerador constata-se que existe uma junção entre a informação fornecida pela função de verosimilhança e a função a priori. Nota-se também que o denominador faz parte da constante de normalização, ficando assim o numerador com toda a informação relevante sobre o parâmetro a estimar.

## 3.2 Função *a priori*

A escolha da função *a priori* é a principal força e fraqueza da estatística Bayesiana. Para os Bayesianos ela permite introduzir informação que pode ser subjetiva, mas que se justifica pelo problema em questão. Para os estatísticos clássicos essa fonte de subjetividade é o problema, uma vez que introduz informação que não pode ser replicada.

Dependendo do problema que se pretende estudar, a função *a priori* pode ser definida/representada como não-informativa ou informativa. A primeira situação surge quando o nosso conhecimento sobre o problema é demasiado limitado e não é possível fazer qualquer tipo de suposição sobre o parâmetro a estimar. Na segunda abordagem é normalmente utilizado conhecimento de dados históricos, de peritos ou uma combinação dos dois para fazer representar o conhecimento através de uma função de densidade.

### 3.2.1 Recolha de Informação

O processo de eliciação é fundamental para a utilização de estatística Bayesiana e acrescenta grande valor aos projetos estatísticos, pois aproxima os trabalhos estatísticos dos seus clientes, incorporando a informação que estes possuem sobre o seu negócio, área de estudo ou problema específico (Garthwaite et al. 2005).

Apesar de extremamente útil, a opinião de peritos pode ser bastante difícil de obter, devido à complexidade envolvida e demora que um bom processo de eliciação requer.

Segundo Garthwaite et al. (2005) o processo de eliciação deve ser composto por 4 fases:

1. Definição dos aspetos do problema que devem ser objeto de eliciação; escolher o perito; treinar o perito para o processo, este passo é crucial caso o perito não esteja familiarizado com probabilidades.
2. Elaboração do método de eliciação, ou seja define-se que perguntas serão feitas ,e de que forma, para obter informação que permita a construção de uma função de distribuição.
3. Representação da informação através de uma distribuição de probabilidade.
4. Avaliação por parte de perito da eliciação. O processo têm de ser iterativo, ou seja, é necessário confirmar que a função representa bem as opiniões do perito, caso contrário será necessário voltar ao segundo passo.

Um dos principais obstáculos para que a fase dois do processo seja bem sucedida são as heurísticas mentais que os seres humanos utilizam quando são questionados sobre probabilidades de um evento. Por isso, este tem sido um campo que também tem sido investigado por psicólogos.

Para realizar a terceira fase, existem diversos métodos e a escolha depende do problema que se está a estudar e do perito.

Um dos primeiros fatores a ter em consideração é se a recolha de informação é feita com um ou mais peritos. O caso de vários peritos não vai ser abordado, uma vez que neste trabalho só será utilizado um . No entanto, é necessário frisar que técnicas de eliciação com o recurso a vários peritos, como o método clássico de Cooke (1991), têm vindo a ser estudadas, principalmente na área de risco operacional, onde o regulador exige que se utilize informação de especialistas para o cálculo dos requisitos de capital. Tal deve-se ao facto de nessa área muito poucos, ou mesmo nenhuns, dados históricos existirem (Bakker 2004).

Recorrendo a um perito existem, ainda assim, várias técnicas que podem ser utilizadas. Neste trabalho vão ser referidas apenas duas. Na primeira assume-se que o conhecimento do perito pode ser representado por uma distribuição de probabilidades específica. Kiefer (2009) usa a função de distribuição beta uma vez que esta está suportada entre 0 e 1, num trabalho ainda muito preliminar sobre a aplicação Bayesiana ao risco de crédito. Nesse caso o perito é inquirido sobre quantis que o autor considerou importantes. O mesmo autor, posteriormente (Jacobs & Kiefer 2010), utiliza uma abordagem diferente utilizando um conceito da física, a Entropia. Neste artigo o autor questiona o perito sobre os quantis da

distribuição da probabilidade de incumprimento. No entanto, uma vez que não quer introduzir suposições sobre a forma da função já definida, decide maximizar a entropia, tendo como restrições os quantis que recebeu. De seguida, para alisar o resultado utiliza o kernel de Epanechnikov.

### 3.3 Monte Carlo via Cadeias de Markov

Um das desvantagens, iniciais, da abordagem Bayesiana perante a Clássica é o facto de, na maior parte dos problemas de interesse, o cálculo da função *a posteriori* de forma analítica e numérica não ser uma opção, principalmente em casos multi-dimensionais. No entanto, a partir da década de 90, devido aos avanços computacionais, passou a ser mais fácil e rápido o cálculo da função *a posteriori* através de métodos de simulação como Monte Carlo e Monte Carlo Via cadeias de Markov (MCMC). É sobre o último método que vai recair a análise nesta secção. Irá ser apresentada uma breve introdução às cadeias de Markov e a métodos de Monte Carlo, pois uma abordagem mais profunda iria consumir muito espaço. O leitor interessado poderá ler sobre estes assuntos em: Tierney (1994), Robert & Casella (2005) e Gamerman & Lopes (2006).

#### 3.3.1 Cadeias de Markov

Nesta secção serão apresentados os conceitos fundamentais sobre cadeias de Markov necessários para se compreender os algoritmos de simulação que serão expostos.

Seja  $\mathbb{X}$  um conjunto,  $\mathcal{F}$  uma  $\sigma$ -álgebra em  $\mathbb{X}$  e  $\pi$  uma distribuição de probabilidade definida em  $(\mathbb{X}, \mathcal{F})$ .

Uma cadeia de Markov  $\{\theta_n\} = \{\theta_n, n = 0, 1, \dots\}$  em  $\mathbb{X}$  diz-se homogênea se a sua função de transição  $P$  tiver as seguintes características:

- Para qualquer conjunto fixo  $A \in \mathcal{F}$  a função  $P(., A)$  é mensurável
- Para qualquer  $x \in \mathbb{X}$ , fixo,  $P(x, .)$  é uma medida de probabilidade em  $(\mathbb{X}, \mathcal{F})$

- para todo o  $n \geq 0$

$$P(\theta_{n+1} \in A | \theta_n = x) = P(x, A)$$

para todo  $x \in \mathbb{X}$  e  $A \in \mathcal{F}$ .

A função transição ao fim de  $i$  passos é dada pela seguinte formula:

$$P(\theta_i \in A | \theta_0 = x) = P^i(x, A) \quad (3.2)$$

O problema que a construção de algoritmos de MCMC induz é o de criar valores correlacionados. Isso levanta o problema do cálculo do valor esperado, uma vez que a lei forte dos grandes números só se aplica a valores independentes. A resposta a este problema será dada no fim da secção, mas primeiro é necessário apresentar algumas definições.

**Definição 1.** A função  $\pi$  é estacionária tendo em conta a função de transição  $P$  se:

$$\pi P(A) = \int P(x, A) \pi(dx) = \pi(A) \quad (3.3)$$



Em MCMC é construída uma cadeia estacionária, no entanto é necessário garantir que ela é única, caso contrário não tem utilidade.

**Definição 2.** A distribuição  $\pi$  é reversível tendo em conta a função de transição  $P$  se:

$$\pi(dx)P(x, dy) = P(y, dx)\pi(dx) \forall x, y \in \mathbb{X} \quad (3.4)$$

A reversibilidade é bastante importante uma vez que se uma cadeia é reversível então é estacionária. É através da reversibilidade que se prova que o algoritmo Metropolis-Hastings gera uma distribuição estacionária.

**Definição 3.** Uma cadeia de Markov é  $\pi$ -irredutível para um distribuição  $\pi$  em  $\mathbb{X}$  se  $\pi(A) > 0$ , com  $A \subset \mathbb{X}$ , implica que:

$$P(\inf\{n \geq 1 : \theta_n \in A\} < \infty | \theta_0 = x) > 0 \quad (3.5)$$

para todo  $x \in \mathbb{X}$

Ou seja, uma cadeia é irredutível se partindo de qualquer estado se consegue chegar a qualquer outro estado.

**Definição 4.** Uma cadeia de Markov  $\theta_n$  é  $\pi$ -irredutível com distribuição estacionária  $\pi$  é recorrente se para cada  $A \subset \mathbb{X}$  com  $\pi(A) > 0$  se verifica que:

$$P(\theta_n \in A \text{ i.o.} | \theta_0 = x) > 0 \text{ para todo o } x, \quad (3.6)$$

$$P(\theta_n \in A \text{ i.o.} | \theta_0 = x) = 1 \text{ para quase todo o } x.$$

O i.o representa infinitas vezes (infinitely often)

**Teorema 1.** *Supondo que a cadeia de Markov  $\theta_n$  é irredutível e tem uma distribuição estacionária  $\pi$ , então a cadeia é  $\pi$ -irredutível,  $\pi$  é a única distribuição estacionária da cadeia e a cadeia é recorrente positiva.*

Já estamos em condições de definir quando a lei forte dos grandes números se aplica às cadeias de Markov.

**Teorema 2.** *Supondo que  $(\theta_n)$  é uma cadeia de Markov irredutível com uma função de transição  $P$  e tem uma distribuição estacionária  $\pi$ . Sendo <sup>1</sup>:*

$$\bar{P}^n(x, A) = \frac{1}{n+1} \sum_{i=0}^n P^i(x, A) \quad (3.7)$$

$\forall x \in A \subset \mathbb{X}$ . Então

$$\|\bar{P}^n(x, \cdot) - \pi(\cdot)\| \rightarrow 0 \quad (3.8)$$

para  $\pi$  em quase todo  $x$ .

**Teorema 3.** *Supondo que  $\theta_n$  é uma cadeia de Markov irredutível com uma função de transição  $K$ , distribuição estacionária  $\pi$  e seja  $f$  uma função real integrável.*

Então:

$$P \left[ \frac{1}{N+1} \sum_{n=0}^N f(\theta_n) \xrightarrow{N \rightarrow \infty} \int f(x) \pi(dx) | \theta_0 = x \right] = 1 \quad (3.9)$$

para quase todos os valores iniciais  $x$ .

Os dois teoremas acima mostram que a proporção observada e esperada do tempo passado no conjunto  $A$  converge para  $\pi(A)$ .

Uma função de transição diz-se periódica se existir um  $d \geq 2$  e uma sequência  $\{E_0, \dots, E_{d-1}\}$  composta por conjuntos não vazios e disjuntos em  $\mathcal{F}$  tal que, para

---

<sup>1</sup>  $\|\cdot\|$  é a norma de variação total

todo  $i = 0, \dots, d - 1$  e para todo  $x \in E_i$ ,

$$P(x, E_i = 1) \text{ para } j = i + 1(\text{mod } d) \quad (3.10)$$

como  $d = 1$  a função de transição diz-se aperiódica (Tierney 1994).

O resultado do teorema 2 pode ser fortalecido caso se considere uma cadeia aperiódica.

**Teorema 4.** *Supondo que  $\theta_n$  é uma cadeia de Markov irreduzível, aperiódica com uma função de transição  $K$  e com uma distribuição estacionária  $\pi$ . Então*

$$\|P^n(x, \cdot) - \pi(\cdot)\| \rightarrow 0 \quad (3.11)$$

para  $\pi$  quase todo  $x$

Com esta pequena introdução às definições e teoremas aplicados a cadeias de Markov tem-se agora condições para a expor os algoritmos de simulação que utilizam cadeias de Markov.

### 3.3.2 Algoritmo Metropolis-Hastings

O algoritmo Metropolis-Hastings faz com que ao gerar números pseudo-aleatórios da distribuição  $q(\cdot|\cdot)$ , normalmente fácil de simular, se corrijam os valores de forma a que estes se aproximem da distribuição objetivo  $\pi$ . Ou seja, já é sabido qual é a distribuição estacionária, assim o problema reside em encontrar a função de transição apropriada.

A ideia base dos algoritmos de cadeias de Markov é a de, através de uma função de transição  $P(x, \cdot)$  gerar uma cadeia que tenha como distribuição estacionária a função que se quer conhecer  $\pi$  (Robert et al. 2010).

O algoritmo tem a seguinte forma:

---

**Algoritmo 1** Metropolis-Hastings

---

```

1: Escolher a função  $q(\cdot|\cdot)$ 
2: Escolher  $\theta_1$ 
3: para  $t \leftarrow 2$  até  $n$  faça
4:   Simular  $Y \sim q(y|\theta_{t-1})$  e  $U \sim U_{[0,1]}$ 
5:   Calcular  $\alpha(\theta_{t-1}, Y) = \min\{\frac{q(\theta_{t-1}|Y)\pi(Y)}{q(Y|\theta_{t-1})\pi(\theta_{t-1})}, 1\}$ 
6:   se  $U \leq \alpha(\theta_{t-1}, Y)$  então
7:      $\theta_t = Y$ 
8:   senão
9:      $\theta_t = \theta_{t-1}$ 
10:  fim se
11: fim para

```

---

Aqui  $\alpha(\theta_{t-1}, Y) = \min\{\frac{q(\theta_{t-1}|Y)\pi(Y)}{q(Y|\theta_{t-1})\pi(\theta_{t-1})}, 1\}$  representa a probabilidade de aceitação do valor proposto. A principal preocupação a ter com o algoritmo é que ele respeite as restrições necessárias para que possa ser usada a lei Forte dos Grandes Números. Para tal é necessário que se consiga provar que através da função proposta se gera uma função estacionária. Neste ponto entra a condição de reversibilidade que é muito útil para provar a estacionariedade. A prova pode ser encontrada em Geyer (2005).

Outra propriedade importante é a irreducibilidade e aperiodicidade da cadeia. Estas podem ser satisfeitas caso a função  $q$  consiga gerar o espaço da função que se quer reproduzir. Ou seja, tem que ter o mesmo suporte.

Por fim, é também necessário que se conheça o rácio  $\frac{\pi(y)}{q(y|x)}$  a menos de uma constante.

O algoritmo Metropolis é um caso particular muito importante. Neste algoritmo utiliza-se uma função  $q(y|x)$  que é simétrica. Essa função proposta simplifica a probabilidade de aceitação, que passa a ser:

$$\alpha(\theta_{t-1}, Y) = \min\left\{\frac{\pi(Y)}{\pi(\theta_{t-1})}, 1\right\} \quad (3.12)$$

A proposta normalmente utilizada é  $y = x + e$  em que  $e$  tem uma distribuição normal de média 0 e variância  $\sigma^2$ . A escolha da variância tem implicações na taxa de aceitação.

Está demonstrado por Tierney (1994) que a função de transição do algoritmo é irreduzível e Harris recorrente, logo gera uma cadeia estacionária única  $\pi$ .

Outro caso importante encontra-se a proposta é independente do valor anterior da cadeia. Esse tipo de algoritmos pode ser muito útil na estatística Bayesiana caso a *função a priori* se assemelhe à distribuição que se quer inferir (Gelman & Lopes 2006).

### 3.3.3 Algoritmo de Gibbs

O algoritmo de Gibbs permite simular distribuições multivariadas e é um caso especial do algoritmo Metropolis-Hastings com uma probabilidade de aceitação igual a 1. O algoritmo gera uma cadeia de Markov através das distribuições condicionais completas ("*full conditionals*") das variáveis. Assim, é requerido, não só, que se conheçam as distribuições condicionais completas, mas que também seja possível gerar números pseudo-aleatório dessas mesmas distribuições (Robert et al. 2010).

Em baixo está representado o algoritmo:

---

#### Algoritmo 2 Gibbs

---

```

1: Escolher  $(\theta_1^1, \dots, \theta_1^J)$ 

2: para  $t \leftarrow 2$  até  $n$  faça

3:   Simular  $\theta_t^1 \sim p_1(\theta_t^1 | \theta_{t-1}^2, \dots, \theta_{t-1}^J)$ 

4:   Simular  $\theta_t^2 \sim p_2(\theta_t^2 | \theta_t^1, \theta_{t-1}^3, \dots, \theta_{t-1}^J)$ 

5:    $\vdots$ 

6:   Simular  $\theta_t^J \sim p_J(\theta_t^J | \theta_t^1, \dots, \theta_{t-1}^{J-1})$ 

7: fim para

```

---

### 3.3.4 Algoritmo de Metropolis-Hastings dentro de Gibbs

Este algoritmo é apenas a inserção do algoritmo de Metropolis-Hastings dentro de cada passo de Gibbs. O algoritmo revela-se de grande utilidade quando é difícil simular as distribuições condicionais completas (Gamerman & Lopes 2006).

---

**Algoritmo 3** Metropolis-Hastings dentro de Gibbs

---

```

1: Escolher  $(\theta_1^1, \dots, \theta_1^J)$ 

2: para  $t \leftarrow 2$  até  $n$  faça

3:   para  $j \leftarrow 1$  até  $J$  faça

4:     Simular  $Y \sim q_j(y|\theta_{t-1}^j)$  e  $U \sim U_{[0,1]}$ 

5:     Calcular  $\alpha(\theta_{t-1}^j, Y) = \min\left\{\frac{q_j(\theta_{t-1}^j|Y)\pi(Y|\theta_{-j})}{q_j(Y|\theta_{t-1}^j)\pi(\theta_{t-1}^j|\theta_{-j})}\right\}$ 

6:     se  $U \leq \alpha(\theta_{t-1}^j, Y)$  então

7:        $\theta_t^j = Y$ 

8:     senão

9:        $\theta_t^j = \theta_{t-1}^j$ 

10:    fim se

11:  fim para

12: fim para

```

---

$$\theta_{-j} = (\theta_t^1, \dots, \theta_t^{j-1}, \theta_{t-1}^{j+1}, \dots, \theta_{t-1}^J)$$

Este método é bastante geral e útil, sem impor mais restrições para além das impostas pelos algoritmos anteriores (Paulino et al. 2003).

Uma análise mais aprofundada sobre este tipo de algoritmos é feita por Chan & Geyer (1994).

### 3.3.5 Análise de resultados das cadeias de Markov

A construção do algoritmo de MCMC é apenas a primeira parte da solução de um problema. De seguida, é necessário analisar de forma cuidada os resultados.

É necessário ter em consideração vários pontos:

1. Se o resultado converge para distribuição estacionária,
2. Se a cadeia percorre todo o espaço da função que se quer estimar,
3. Ao contrário dos métodos de Monte Carlo, os algoritmos de MCMC criam amostras correlacionadas.

Para a análise do diagnóstico da convergência pode-se utilizar o método de Geweke et al. (1991) que tem como base o teste da diferença das médias. O método utiliza os primeiros 10% dos valores da cadeia e os últimos 50%. O que se pretende investigar é se as duas partes da cadeia advêm da mesma distribuição (Hipótese nula). A estatística de Geweke é um Z-score normalizado em que os desvios padrão são ajustados para a auto-correlação.

Raftery & Lewis (1996) apresentaram um método para encontrar o número de valores a serem descartados do início da cadeia ("*burn-in*"). Este método consiste em estimar o quantil  $q$  com um erro de  $+/-r$  com uma probabilidade de  $pro$ . O  $q$ ,  $r$  e  $pro$ , normalmente utilizados são, .025, .005 e .95 respectivamente.

Quanto ao segundo ponto, nunca se pode garantir que a simulação visita todo o espaço, no entanto existem técnicas para minimizar esse erro.

Uma regra utilizada na literatura é a escolha de distribuições propostas que gerem uma taxa de aceitação perto de 25% para problemas com muitos parâmetros a estimar e 50% para problemas com um ou dois parâmetros Robert et al. (2010).

Por fim, pode-se optar pelo método "*batch means*" utilizado por Geyer (1992), para tratar a correlação. O método consiste em separar a amostra em  $b$  amostras e fazer a média das mesmas e utilizar o erro standard de Monte Carlo (MCSE) como critério de paragem. Ou seja, quanto mais elevado o MCSE me-



nos confiável é o resultado. Sendo  $N$  o número de simulações e  $b$  a quantidade de números em cada "batch":

$$\hat{\mu}_{b,k} = \frac{1}{b} \sum_{i=bk+1}^{bk+b} g(X_i) \quad (3.13)$$

em que  $k = 1, \dots, N/b$ .

A escolha de  $b$  advém de uma análise à correlação da cadeia. Escolhe-se então um valor  $b$  que faça com que não exista correlação entre diferentes  $\hat{\mu}_{b,k}$ . A vantagem deste método é a capacidade de, utilizando  $\mu_{b,k}$ , se poder aplicar o  $t$  – teste e obter um intervalo de confiança da média. Como cada  $\hat{\mu}_{b,k}$  é aproximada por uma distribuição normal de média  $\mu$  e variância  $\sigma^2/b$ . A a variância empírica é:

$$\sigma^2 = \frac{b}{N} \sum_{k=1}^{N/b} (\mu_{b,k} - \hat{\mu}_N)^2 \quad (3.14)$$

O MCSE é  $\frac{\sigma}{\sqrt{N}}$ .

## 4. Metodologia e Resultados

Neste capítulo serão analisados, em primeiro lugar, os dados. De seguida, é necessário definir um modelo económico a utilizar para representar a função de verosimilhança. Na terceira fase constrói-se a distribuição *a priori*, escolhendo o especialista e definindo, neste caso, os quantis a eliciar. Tendo a informação do perito representada passa-se para a construção da função *a posteriori* através de cadeias de Markov. A implementação do algoritmo foi feita através da combinação das linguagens R e C++. Por fim, é realizada uma análise dos resultados da simulação utilizando as metodologias descritas em 3.3.5.

### 4.1 Dados

Neste estudo foi utilizada uma carteira que contém grandes empresas portuguesas. Em baixo é feita a descrição dos dados:

- Período Histórico: Setembro de 2010 até Setembro de 2013. Ou seja dados de 4 anos (por conveniência utiliza-se o mês de setembro como final).
- A carteira é composta por 2032 empresas.
- Os empréstimos são classificados através de uma escala de rating de 1 a 100, sendo 1 o rating mais baixo e 100 o mais elevado (o que apresenta menos risco).
- Os empréstimos têm um acompanhamento mensal da sua situação, em

que 1 representa incumprimento e 0 uma situação regular.

Os dados possuem alguns problemas evidenciados através de uma análise detalhada que justificam a escolha de uma abordagem Bayesiana. Por exemplo existem empréstimos com falta de acompanhamento da sua situação (regular ou em incumprimento). Ou seja, existem meses em que não há qualquer tipo de informação sobre se o crédito esteve regular ou não. Por outro lado, o histórico da carteira é demasiado curto, apenas 4 anos.

Assim, apesar de os dados históricos continuarem ser importantes justifica-se a introdução de informação de um perito na análise do risco da carteira.

Na tabela abaixo pode-se observar algumas características da carteira.

Rating	$\leq 50$	$> 50$
Nº Total	1319	713
Contratos com 46 meses	911	477

Tabela 4.1: Dados da carteira Grandes Empresas

O estudo irá focar-se nas grandes empresas com um rating superior a 50, uma vez que a abordagem Bayesiana é mais relevante precisamente para carteiras com poucos dados de incumprimento. Assim, a carteira final tem 477 empresas.

Os dados sobre o incumprimento são os seguintes: em 2010 não existiram contratos em incumprimento; em 2011 existiu 1; em 2012 houve 10; em 2013 observaram-se 18.

## 4.2 Modelo

O modelo escolhido para representar a informação de dados históricos é uma extensão do modelo estrutural de um único fator de Vasicek.

A extensão ao modelo original de Vasicek é feita através da introdução de uma análise a vários períodos de tempo que estão correlacionados através do risco sistémico.

O modelo pode ser visto em Lamb & Perraudin (2008) e Jacobs & Kiefer (2010), com poucas diferenças. A única diferença está na representação da correlação temporal. Apesar de o modelo que irá ser utilizado ser o de Lamb & Perraudin (2008), para uma melhor compreensão da intuição económica do modelo deve-se ler Jacobs & Kiefer (2010), uma vez que apresenta gradualmente a derivação do modelo e o artigo tem uma como ambição ser um guia para a utilização de estatística bayesiana na modelação do risco de crédito.

O modelo assume que numa carteira diversificada o risco idiossincrático  $e_t$  é neutralizado e só o risco sistémico  $z_t$  afeta o valor da empresa.

Supondo que o valor de uma empresa ( $v_t$ ), no momento  $t$ , é representado por:

$$\begin{aligned} v_t &= \sqrt{\rho} z_t + \sqrt{1 - \rho} e_t \\ z_t &= \tau z_{t-1} + \sqrt{1 - \tau^2} w_t \end{aligned} \tag{4.1}$$

onde  $t = 1, \dots, T$  e  $|\tau| < 1$ .

A correlação temporal ( $\tau$ ), feita através do fator de risco sistémico, respeita o facto de o valor da empresa ser representado por uma distribuição normal standard.

A variável  $\rho$  é interpretada como a correlação entre o valor da empresa e o risco sistêmico. As variáveis  $e_1, \dots, e_T$  e  $w_1, \dots, w_T$  são i.i.d.  $N(0, 1)$ .

Normalmente a variável  $z_t$  ("o shock da economia") tem variância unitária. O modelo assim respeita essa suposição, uma vez que a distribuição não condicional de  $z_t$  tem variância 1. Esta é a diferença entre o modelo representado em Lamb & Perraudin (2008) e Jacobs & Kiefer (2010). A escolha do modelo autoregressivo de Jacobs & Kiefer (2010) não respeitava esta condição.

Sendo um modelo estrutural, o incumprimento acontece quando o valor da empresa é inferior a um determinado limite  $c$ .

Sendo  $Y$  uma função indicadora do incumprimento, então:

$$\begin{aligned} Y_t &= I(v_t < c) \\ p &= P(Y_t = 1) = P(v_t < c) = \Phi(c) \end{aligned} \tag{4.2}$$

A probabilidade de incumprimento é representada por:

$$\begin{aligned} g_t(p, \rho, z_t) &= P(Y_t = 1 | p, \rho, z_t) \\ &= P(v_t < c | p, \rho, z_t) \\ &= \Phi\left(\frac{\Phi^{-1}(p) - \sqrt{\rho}z_t}{\sqrt{1 - \rho}}\right) \end{aligned} \tag{4.3}$$

No modelo existe também a simplificação de que quando se está a estudar uma carteira em que os empréstimos possuem o mesmo risco, carteira homogênea, o número de incumprimentos pode ser modelado por uma distribuição binomial.

Assim, sendo  $\mathbf{y} = (y_1, \dots, y_T)$  o número de incumprimentos observados e  $\mathbf{n} =$

$\{n_1, n_2, \dots, n_T\}$  o tamanho da carteira, a função de densidade de  $y_t$  dado  $z_t, p, \rho$  é:

$$f(y_t|p, \rho, z_t) = \binom{y_t}{n_t} g_t(p, \rho, z_t)^{y_t} (1 - g_t(p, \rho, z_t))^{n_t - y_t} \quad (4.4)$$

onde  $\mathbf{y} = (y_1, \dots, y_T)$  e  $\mathbf{z} = (z_1, \dots, z_T)$ .

Consequentemente a função de verosimilhança que irá ser utilizada na função *a posteriori* é dada por:

$$f(\mathbf{y}|p, \rho, \mathbf{z}) = \prod_{t=1}^T \binom{y_t}{n_t} g_t(p, \rho, z_t)^{y_t} (1 - g_t(p, \rho, z_t))^{n_t - y_t} \quad (4.5)$$

#### 4.2.1 Função *a priori*

Tendo o modelo várias variáveis,  $p, \rho, \tau, \mathbf{z}$  é necessário introduzir funções *a priori* para cada uma. Para a função de  $p$  decidiu-se recorrer a um perito. A abordagem utilizada foi semelhante à de Kiefer (2009). O que se pretende eliciar é então a função de probabilidade da probabilidade de incumprimento tendo como horizonte temporal 1 ano. O perito está bastante familiarizado com estatística, o que facilitou muito o processo.

Foram inquiridos os quantis: 25%, 50%, 75%, 90%, 95%. A resposta obtida foi: 1.3%, 2%, 3%, 4.5%, 7%.

De seguida, modelou-se a função assumindo que ela correspondia a uma distribuição beta. A modelação é feita através da minimização da soma dos erros quadráticos entre os quantis da distribuição beta e os quantis fornecidos. Para tal foi utilizado a função "get.beta.par" do pacote "rrsikDistributions". O resultado foi uma distribuição beta com  $\alpha = 2.5$  e  $\beta = 104.98$ , ou seja uma distribuição beta

com valor esperado  $\approx 0.0233$  e variância  $\approx 0.000209$ .

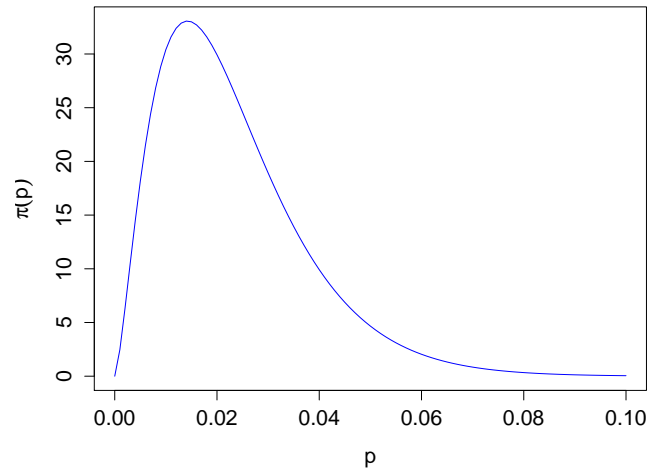


Figura 4.1: Distribuição *a priori*

No anexo B pode ver-se que o desfasamento entre os quantis da distribuição beta definida e os quantis do perito não é significativo.

Quanto ao parâmetro  $\rho$ , seria muito difícil conseguir utilizar a informação do especialista, uma vez que tem muito pouca sensibilidade para conseguir definir uma função para a variável. Decidiu-se optar por "*eliminar a variável*" utilizando a função definida por Kiefer (2011):

$$\rho(p) = .24 - .12(1 - e^{-50p}) \quad (4.6)$$

Esta fórmula advém da eliminação do fator  $\frac{1}{1-e^{-50}}$  da fórmula definida em Basileia II, uma vez que apenas difere em 1 na 22<sup>a</sup> casa decimal.

Na base da fórmula estão factos empíricos, bem como intuição económica. Ela supõe que quanto maior a PD menor a correlação com o fator sistémico. A razão é simples, se a empresa tem uma probabilidade de incumprimento elevada

significa que está mais exposta ao seu próprio risco do que ao risco sistémico (Basel Committe on Banking Supervision 2005a).

Devido ao facto de existir muito pouca informação sobre o parâmetro  $\tau$ , optou-se por utilizar o princípio da razão insuficiente de Laplace escolhendo a distribuição  $U_{]-1,1[}$ , ou seja uma distribuição não-informativa.

## 4.2.2 Função *a posteriori*

A função *a posteriori* é a seguinte:

$$\pi(p, \mathbf{z}, \tau | \mathbf{y}) = \frac{f(\mathbf{y} | p, \mathbf{z}) \pi(p) \pi(\mathbf{z} | \tau) \pi(\tau)}{f(\mathbf{y})} \quad (4.7)$$

onde  $\pi$  representa a função *a priori*. Sendo as variáveis independentes têm-se as seguintes distribuições condicionais completas:

Função condicionada da probabilidade de incumprimento:

$$\begin{aligned} \pi(p | \mathbf{y}, \mathbf{z}, \tau) &\propto f(\mathbf{y} | p, \mathbf{z}) \pi(p) \\ &\propto \left[ \prod_{t=1}^T g_t(p, z_t)^{y_t} (1 - g_t(p, z_t))^{n_t - y_t} \right] p^\alpha (1 - p)^{\beta - 1} \end{aligned} \quad (4.8)$$

Função condicionada da correlação temporal  $\tau$ :

$$\begin{aligned} \pi(\tau | \mathbf{y}, \mathbf{z}, p) &\propto f(\mathbf{z} | \tau) \pi(\tau) \\ &\propto (1 - \tau^2)^{-\frac{T-1}{2}} \exp\left(-\sum_{t=2}^T \frac{(z_t - \tau z_{t-1})^2}{2(1 - \tau^2)}\right) \end{aligned} \quad (4.9)$$



Função condicionada do fator sistêmico no momento 1:

$$\begin{aligned}\pi(z_1|\mathbf{y}, \mathbf{z}_{-1}, \tau) &\propto f(\mathbf{y}|p, z_1)\pi(z_1) \\ &\propto g_1(p, z_1)^{y_1}(1 - g_1(p, z_1))^{n_1 - y_1} \exp\left(\frac{-z_1^2}{2}\right)\end{aligned}\tag{4.10}$$

Função condicionada do fator sistêmico no momento  $> 1$ :

$$\begin{aligned}\pi(z_t|\mathbf{y}, \mathbf{z}_{-t}, \tau) &\propto f(\mathbf{y}|p, \mathbf{z})\pi(z_t|z_{t-1}, \tau) \\ &\propto g_t(p, z_t)^{y_t}(1 - g_t(p, z_t))^{n_t - y_t} \exp\left(-\frac{(z_t - \tau z_{t-1})^2}{2(1 - \tau^2)}\right)\end{aligned}\tag{4.11}$$

## 4.2.3 Implementação e Resultados

Para a implementação do modelo desenvolveu-se um algoritmo de MCMC. O algoritmo utilizado é o algoritmo de Metropolis-Hastings dentro de Gibbs. Esse algoritmo foi desenvolvido através da combinação da linguagem de programação R e C++. O trabalho foi sempre desenvolvido no ambiente de desenvolvimento integrado (IDE) RStudio.

A combinação do R com o C++ apresenta grandes vantagens e é feita através do pacote RCPP<sup>1</sup> e RCPPArmadillo<sup>2</sup>. A principal razão para a combinação das linguagens é aproveitar a rapidez do C++, uma vez que se irá simular uma grande quantidade de números pseudo-aleatórios, e ao mesmo tempo ser prático criar gráficos e utilizar pacotes de diagnóstico de MCMC, como o CODA. O algoritmo pode ser visto no anexo C.

Uma etapa crucial em MCMC é a escolha das distribuições propostas.

<sup>1</sup>Faz a ligação entre o C++ e o R

<sup>2</sup>Biblioteca em C++ de Álgebra Linear

Para  $p(\theta_n^1)$  foi escolhida uma distribuição beta que toma como valor esperado o valor anterior da cadeia:  $\theta_n^1 | \theta_{n-1}^1 \sim B(c\theta_{n-1}^1, c(1-\theta_{n-1}^1))$ . O valor  $c$  e  $\theta^1$ -inicial foram escolhidos para que a distribuição inicial fosse a distribuição *a priori* ( $c = 107$  e  $\theta_0^1 = 0.023$ ).

A escolha da função beta é adequada uma vez que tem o mesmo suporte de  $p$ ,  $[0, 1]$ .

Para a variável  $\tau$  escolheu-se a distribuição  $U_{[-1,1]}$ . A escolha deve-se ao facto de não existir qualquer informação sobre o seu valor. Assim, neste passo irá ser usado um algoritmo de Metropolis-Hastings independente.

As variáveis  $z_t$  foram simuladas através do algoritmo de Metropolis utilizando uma distribuição  $N(0, 1)$  para definir o salto.

Foi gerada uma cadeia de 1e6 valores aleatórios e um "burn-in" de 1000. No anexo D.1 pode-se observar os testes aplicados ao resultado da simulação de Markov.

No anexo D.1.2 é apresentado o teste de Raftery que mostra que o "burn-in" utilizado é superior ao necessário, no entanto é utilizado por uma questão de aumentar a segurança.

O tamanho de cada "batch" foi de 250. No anexo D.2 pode-se observar a auto-correlação.

Quanto à convergência, não existe qualquer evidência que aponte para que esta não tenha sido atingida (ver anexo D.1).

As taxas de aceitação foram as seguintes:  $(\theta^1, \theta^2, \dots, \theta^6) = (.262, .504, .565, .505, .331, .278)$

De seguida, passa-se para a utilização do método "batch means". A análise

realizada é bastante semelhante à feita por (Geyer 2012).

O valor esperado da PD e  $\tau$ , bem como o MCSE são dados na seguinte tabela:

	$p$	$\tau$	$z_1$	$z_2$	$z_3$	$z_4$
Estimação	0.01994	0.26851	1.47956	1.09970	-0.42610	-1.03491
MCSE	6.851e-05	7.801e-04	3.114e-03	3.752e-03	3.944e-03	3.609e-03

Tabela 4.2: Médias das variáveis *a posteriori*

Quanto à forma da função da densidade da PD é a seguinte:

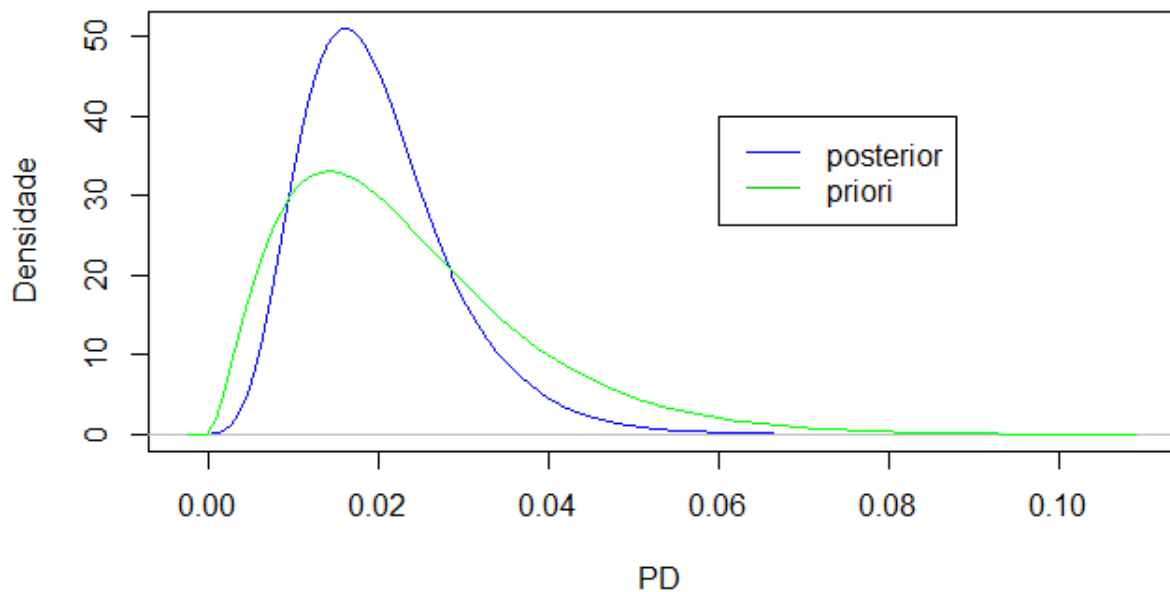


Figura 4.2: Função  $\pi(p|\mathbf{y})$

Ao observar a figura 4.2 pode-se tirar a conclusão de que apesar de o valor esperado de incumprimento ser bastante semelhante ao definido pelo perito, a função *a posteriori* é mais informativa, uma vez que se encontra mais concen-

trada. Diminui-se, portanto, a dispersão de valores da probabilidade de incumprimento.

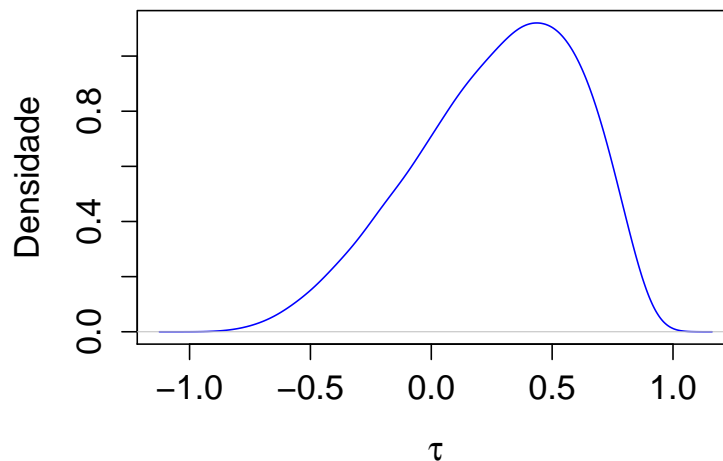


Figura 4.3: Função  $\pi(\tau|\mathbf{y})$

Quanto à correlação temporal observa-se que é positiva, como seria de esperar. Ou seja, a períodos em que o risco sistémico é mais elevado seguem-se períodos de elevado risco sistémico.

A tabela 4.2 mostra que o fator sistémico é positivo quando não existem incumprimentos e à medida que os incumprimentos aumentam o fator sistémico piora, como seria de esperar.

Por fim, definida a função *a posteriori* da PD da carteira de empréstimos seria possível calcular o capital económico necessário, utilizando a LGD definida por Basileia para este tipo de carteiras. No entanto, essa análise fica fora do âmbito do projeto.

## 5. Conclusões

Este trabalho teve como principal objetivo calcular a probabilidade de incumprimento através da incorporação da informação histórica, de uma carteira de crédito real, com informação fornecida por um perito de análise de risco de crédito.

Ao longo do projeto ficou explicitado a importância da introdução da opinião de peritos, principalmente quando a informação histórica é insuficiente ou pouco fiável.

No entanto, a utilização de apenas um perito pode ter desvantagens principalmente se existirem interesses pessoais em causa. Por isso recomenda-se que se desenvolva uma metodologia com um maior número de peritos, caso se de-seje aplicar este modelo na realidade. Para tal, recomenda-se o modelo clássico de Cooke (1991).

Foi também desenvolvido um algoritmo de simulação para determinar a *função a posteriori*. É de facto crucial o conhecimento de Monte Carlo via cadeias de Markov quando se estuda estatística Bayesiana.

Seria também interessante realizar um estudo comparativo, de médio prazo, entre a abordagem aqui apresentada e a abordagem utilizada pelos bancos de forma a perceber qual a mais adequada.

No que toca à correlação temporal do fator económico denota-se que existe uma correlação positiva, apesar de não ser muito forte ( $\approx .269$ ).

Por fim, recomenda-se que se estude a aplicação de estatística Bayesiana a

modelos que combinem o cálculo da LGD com PD. Já existe, pelo menos, um estudo sobre o tema (Shevchenko & Luo 2012), no entanto ainda é uma área a investigar.

# Referências Bibliográficas

Antão, P. & Lacerda, A. (2008), 'Avaliação dos requisitos de capital sob basileia ii: o caso português', *Relatório de Estabilidade Financeira 2008–Banco de Portugal* pp. 207–28.

Bakker, M. (2004), *Quantifying operational risk within banks according to Basel II*, Bakker, MRA.

Banco de Portugal (2014), 'Sistema bancário português - desenvolvimentos recentes (1º trimestre 2014)'.

Basel Committee on Banking Supervision (2000), Principles for the management of credit risk, Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2005a), 'An explanatory note on the basel ii irb risk weight functions'.

Basel Committee on Banking Supervision (2005b), 'Validation of low-default portfolios in the basel ii framework'.

Basel Committee on Banking Supervision (2006), 'International convergence of capital measurement and capital standards'.

- Chan, K. S. & Geyer, C. J. (1994), 'Discussion: Markov chains for exploring posterior distributions', *The Annals of Statistics* pp. 1747–1758.
- Cooke, R. (1991), *Experts in uncertainty: opinion and subjective probability in science*, Oxford University Press, New York.
- Gamerman, D. & Lopes, H. F. (2006), *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*, CRC Press.
- Garthwaite, P. H., Kadane, J. B. & O'Hagan, A. (2005), 'Statistical methods for eliciting probability distributions', *Journal of the American Statistical Association* **100**(470), 680–701.
- Geweke, J. et al. (1991), *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, Vol. 196, Federal Reserve Bank of Minneapolis, Research Department.
- Geyer, C. J. (1992), 'Practical markov chain monte carlo', *Statistical Science* pp. 473–483.
- Geyer, C. J. (2005), 'Markov chain monte carlo lecture notes'. Disponível em: <http://www.stat.umn.edu/geyer/f05/8931/n1998.pdf> [Acesso em : 2014/3/21].
- Geyer, C. J. (2012), 'Markov chain monte carlo and bayesian inference'. Disponível em: <http://www.stat.umn.edu/geyer/s12/5102/notes/mcmc.pdf> [Acesso em : 2014/3/21].



- Jacobs, M. & Kiefer, N. M. (2010), The bayesian approach to default risk: A guide, Technical report, CAE Working Paper.
- Kiefer, N. M. (2009), 'Default estimation for low-default portfolios', *Journal of Empirical Finance* **16**(1), 164–173.
- Kiefer, N. M. (2011), 'Default estimation, correlated defaults, and expert information', *Journal of Applied Econometrics* **26**(2), 173–192.
- Lamb, R. & Perraudin, W. (2008), 'Dynamic default rates'. Disponível em: [http://www.riskcontrollimited.com/research\\_papers/Lamb\\_Perraudin\\_ref8\\_4.pdf](http://www.riskcontrollimited.com/research_papers/Lamb_Perraudin_ref8_4.pdf) [Acesso em : 2014/3/10].
- Paulino, C. D. M., Turkman, M. A. A. & Murteira, B. (2003), *Estatística bayesiana*, Serviço de Educação e Bolsas, Fundação Calouste Gulbenkian, Lisboa.
- Raftery, A. E. & Lewis, S. M. (1996), Implementing mcmc, in 'Markov chain Monte Carlo in practice', Springer, pp. 115–130.
- Robert, C. P., Casella, . o. et al. (2010), *Introducing Monte Carlo Methods with R*, Vol. 18, Springer.
- Robert, C. P. & Casella, G. (2005), *Monte Carlo statistical methods*, Vol. 319, 2 edn, Springer.
- Shevchenko, P. V. & Luo, X. (2012), 'Dependent default and recovery: Markov chain monte carlo study of downturn loss given default credit risk model', *AN-ZIAM Journal* **53**, C185–C202.

Tierney, L. (1994), 'Markov chains for exploring posterior distributions', *the Annals of Statistics* pp. 1701–1728.

## A. Esquema Clássico e Bayesiano

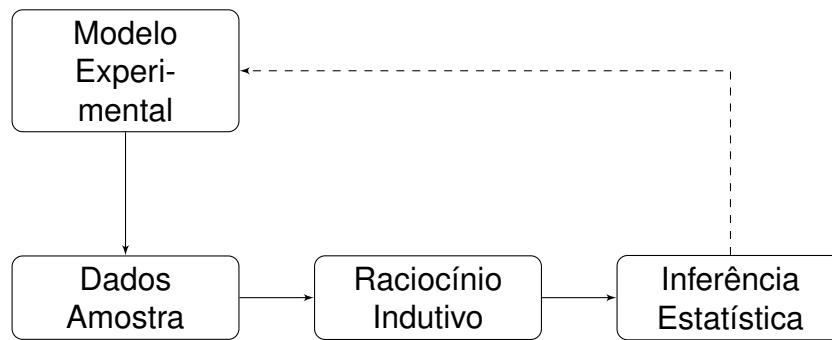


Figura A.1: Esquema Clássico (Paulino et al. 2003)

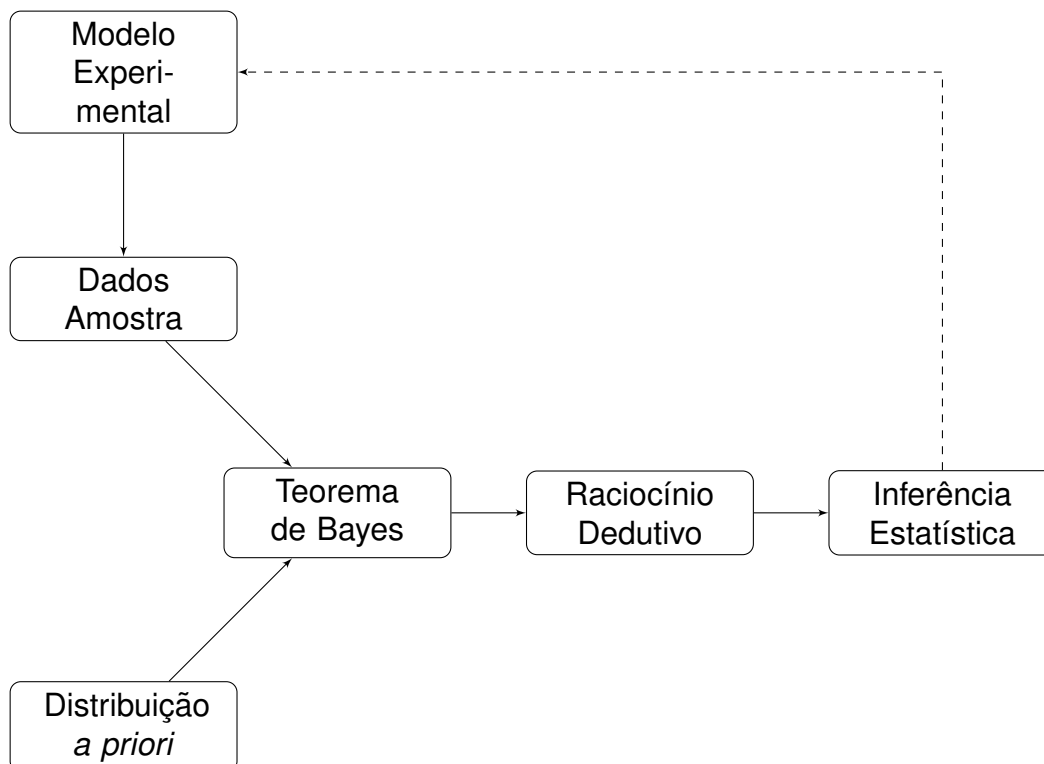


Figura A.2: Esquema Bayesiano (Paulino et al. 2003)

## B. Distribuição *a priori*

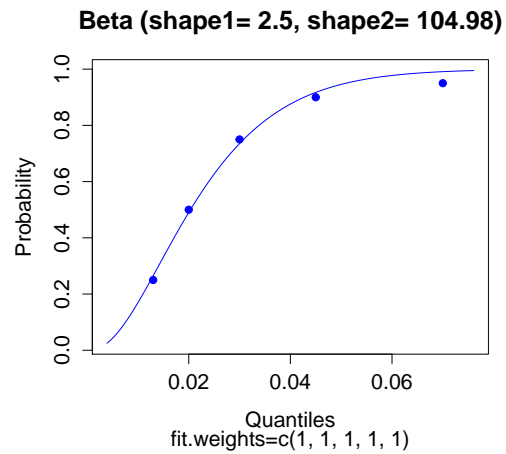


Figura B.1: Desfasamento dos quantis da distribuição *a priori* beta e dos quantis do especialista

## C. Algoritmo M-H dentro de Gibbs

---

**Algoritmo 4** Metropolis-Hastings dentro de Gibbs Utilizado

---

```
1: Escolher  $(\theta_1^1, \theta_1^2, \dots, \theta_1^6)$ 
2: para  $t \leftarrow 2$  até  $n$  faça
3:   Simular  $Y \sim B(c(\theta_{t-1}^1), c(1 - \theta_{t-1}^1))$  e  $U \sim U_{[0,1]}$ 
4:   Calcular  $\alpha(\theta_{t-1}^1, Y) = \min\{\frac{q(\theta_{t-1}^1)\pi(Y|\theta_{t-1}^1)}{q(Y)\pi(\theta_{t-1}^1|\theta_{t-1}^1)}, 1\}$ 
5:   se  $U \leq \alpha(\theta_{t-1}^1, Y)$  então
6:      $\theta_t^1 = Y$ 
7:   senão
8:      $\theta_t^1 = \theta_{t-1}^1$ 
9:   fim se
10:  Simular  $Y \sim U_{[-1,1]}$  e  $U \sim U_{[0,1]}$ 
11:  Calcular  $\alpha(\theta_{t-1}^2, Y) = \min\{\frac{\pi(Y|\theta_{t-1}^2)}{\pi(\theta_{t-1}^2|\theta_{t-1}^2)}, 1\}$ 
12:  se  $U \leq \alpha(\theta_{t-1}^2, Y)$  então
13:     $\theta_t^2 = Y$ 
14:  senão
15:     $\theta_t^2 = \theta_{t-1}^2$ 
16:  fim se
17:  para  $j \leftarrow 3$  até  $6$  faça
18:    Simular  $Y \sim \theta_{t-1}^j + N(0, 1)$  e  $U \sim U_{[0,1]}$ 
19:    Calcular  $\alpha(\theta_{t-1}^j, Y) = \min\{\frac{\pi(Y|\theta_{t-1}^j)}{\pi(\theta_{t-1}^j|\theta_{t-1}^j)}, 1\}$ 
20:    se  $U \leq \alpha(\theta_{t-1}^j, Y)$  então
21:       $\theta_t^j = Y$ 
22:    senão
23:       $\theta_t^j = \theta_{t-1}^j$ 
24:    fim se
25:  fim para
26: fim para
```

---

$$\theta_{-j} = (\theta_t^1, \dots, \theta_t^{j-1}, \theta_t^{j+1}, \dots, \theta_t^J)$$

# D. Diagnóstico Metropolis-Hastings

## D.1 Convergência

### D.1.1 Médias e Teste Geweke

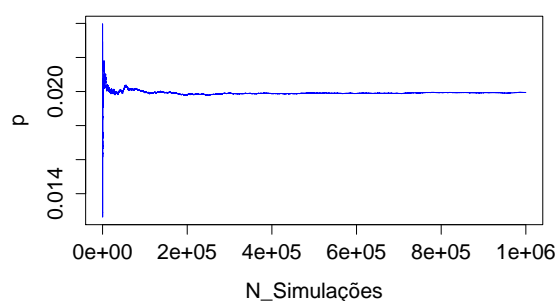


Figura D.1: Evolução da Média da variável  $p$

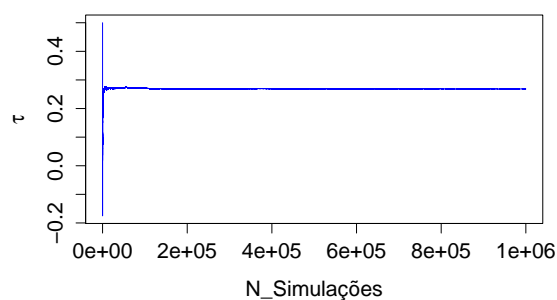


Figura D.2: Evolução da Média da variável  $\tau$

	$p$	$\tau$	$z_1$	$z_2$	$z_3$	$z_4$
Resultados	-0.154	1.045	0.055	-0.363	-.264	-.462

Tabela D.1: Teste Geweke

Não existe nenhum valor superior a 1.96. Ou seja, para um intervalo de

confinação de 95% não se rejeita a hipótese nula (as médias são iguais).

### D.1.2 Teste Raftery e Lewis

	Burn-in	Total	Lower bound	Dependence factor
$p$	140	170324	3746	45.50
$\tau$	8	15888	3746	4.24
$z_1$	51	69615	3746	18.60
$z_2$	66	89738	3746	24.00
$z_3$	124	141856	3746	37.90
$z_4$	124	154473	3746	41.20

Tabela D.2: Teste Raftery e Lewis

## D.2 Correlação

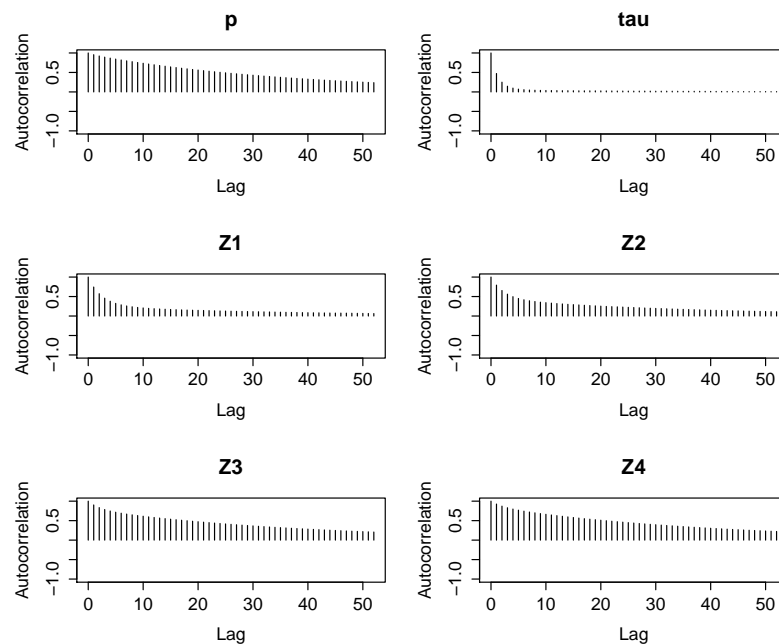


Figura D.3: Correlação dos resultados

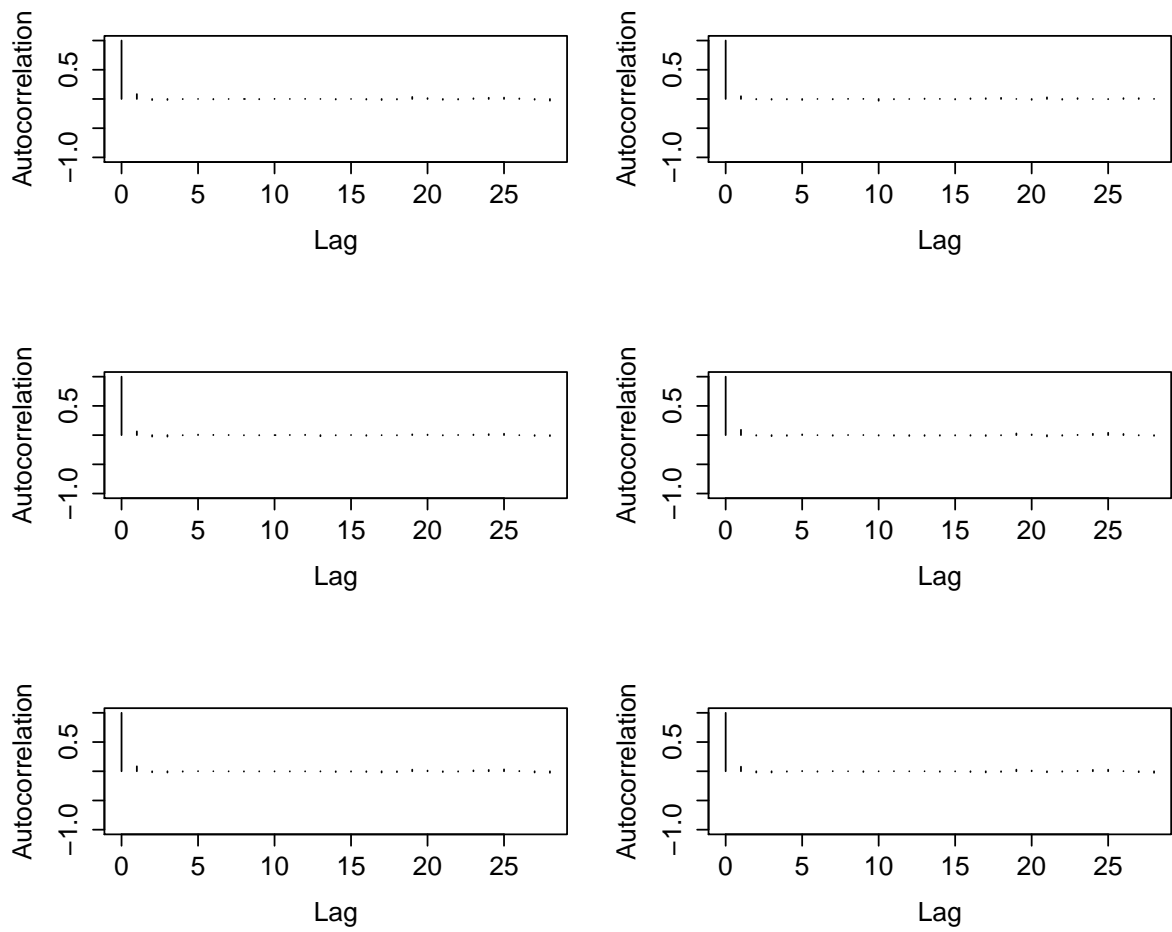


Figura D.4: Correlação depois do método Batch